

This is the peer reviewed version of the following article:

Design, synthesis, biological evaluation and crystal structure determination of dual modulators of carbonic anhydrases and estrogen receptors / Tinivella, Annachiara; Nwachukwu, J. C.; Angeli, A.; Foschi, F.; Benatti, Annalaura; Pinzi, Luca; Izard, T.; Ferraroni, M.; Erumbi, R.; Christodoulou, M. S.; Passarella, D.; Supuran, C.; Nettles, K. W.; Rastelli, Giulio. - In: EUROPEAN JOURNAL OF MEDICINAL CHEMISTRY. - ISSN 1768-3254. - 246:(2023), pp. 115011-115020. [10.1016/j.ejmech.2022.115011]

Terms of use:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

15/08/2024 02:44

(Article begins on next page)

Design, Synthesis, Biological evaluation and Crystal Structure Determination of Dual Modulators of Carbonic Anhydrase and Estrogen Receptor

Annachiara Tinivella,^{1,2} Jerome C. Nwachukwu,³ Andrea Angeli,⁴ Francesca Foschi,^{1,5} Anna Laura Benatti,¹ Luca Pinzi,¹ Tina Izard,³ Marta Ferraroni,⁶ Rangarajan Erumbi,³ Michael S. Christodoulou,^{1,5,7} Daniele Passarella,⁵ Claudiu Supuran,⁴ Kendall W. Nettles,³ Giulio Rastelli^{1*}

¹Department of Life Sciences, University of Modena and Reggio Emilia, Via Giuseppe Campi 103, 41125 Modena, Italy.

² Clinical and Experimental Medicine PhD Program, University of Modena and Reggio Emilia, Modena, Italy.

³ Department of Integrative Structural and Computational Biology, University of Florida Scripps Biomedical Research, 130 Scripps Way, Jupiter, Florida, 33458, USA.

⁴ NEUROFARBA Department, Sezione di Scienze Farmaceutiche, University of Florence, Via Ugo Schiff 6, 50019 Sesto Fiorentino, Florence, Italy.

⁵ Department of Chemistry, University of Milano. Via Golgi 19, 20133 Milano, Italy.

⁶ Department of Chemistry “Ugo Schiff”, University of Florence, Via della Lastruccia 13, 50019 Sesto Fiorentino, Florence, Italy.

⁷ Present address: Department of Food, Environmental and Nutritional Sciences (DeFENS), University of Milan, via Celoria 2, 20133 Milan, Italy.

* **Correspondence to:** Prof. Giulio Rastelli, Department of Life Sciences, University of Modena and Reggio Emilia, Via Giuseppe Campi 103, 41125 Modena, Italy. Tel +39 059 2058564, Email giulio.rastelli@unimore.it

ABSTRACT

Multi-target compounds have become increasingly important for the development of safer and more effective drug candidates. In this work, we devised a combined ligand-based and structure-based multi-target repurposing strategy, and applied it to a series of hexahydrocyclopenta[*c*]quinoline compounds synthesized previously. The *in silico* analyses identified Carbonic Anhydrases (*hCA*) and Estrogen Receptors (ER) as top scoring candidates for dual modulation. Carbonic Anhydrase (*hCA*) isoforms and Estrogen Receptor (ER) subtypes ER α and/or ER β , are co-expressed in various cancer cell types, including breast and prostate cancer cells. ER α is the primary target of anti-estrogen therapy in breast cancer, and the *hCA* IX isoform is a therapeutic target in triple-negative breast cancer. ER α -mediated transcriptional programs and *hCA* activity in cancer cells promote favorable microenvironments for cell proliferation. Interestingly, several lines of evidence indicate that combined modulation of these two targets may provide therapeutic benefits. Moving from these first results, two additional hexahydrocyclopenta[*c*]quinoline derivatives bearing a sulfonamide zinc binding group (*hCA*) and a phenolic hydroxyl (ER) pharmacophoric groups placed at the appropriate locations were designed and synthesized. Interestingly, these compounds were able to directly modulate the activities of both *hCA* and ER targets. In cell-based assays, they inhibited proliferation of breast and prostate cancer cells with micro-molar potency and cell type-selective efficacy. The compounds inhibited *hCA* activity with nano-molar potency and isoform-selectivity. In transactivation assays, they reduced estrogen-driven ER activity with micro-molar potency. Finally, crystal structures of the synthesized ligands in complex with the two targets revealed that the compounds bind directly to the *hCA* active site as well as to the ER ligand-binding domain, providing structural explanation to the observed activity and a rationale for optimization of their dual activity. To the best of our knowledge, this work describes the design, synthesis and biological characterization of the first dual modulators of *hCA* and ER, laying the ground for the structure-based optimization of their dual activity.

KEYWORDS: Carbonic anhydrases; Estrogen Receptor; Cancer; Polypharmacology; Drug repurposing, Virtual Screening; Drug design

INTRODUCTION

Multi-target drugs may provide several potential advantages over single target drugs or combination therapies, including higher efficacy, fewer side effects, less drug resistance and lack of drug-drug interactions.[ref Anighoro JMC 2014]

In this study, we devised and applied an integrated ligand- and structure-based multi-target repurposing approach in search of new potential therapeutic targets of a series of hexahydrocyclopenta[*c*]quinoline compounds synthesized previously in our group [REF]. The available library of compounds was firstly screened for ligand-based similarity with respect to compounds reported in the DrugBank, Protein Data Bank (PDB) and ChEMBL databases [ref], leading to the identification of Carbonic Anhydrases (*hCAs*) and Estrogen Receptor (ER) α and β as putative targets for some of the investigated compounds.

Carbonic anhydrases participate in the regulation of several cellular processes [REF]. Inhibitors of cytosolic isoforms of this protein family have been investigated, for example, for the treatment of retinal and cerebral edema (*hCA I*)[ref], glaucoma, edema and epilepsy (*hCA II*, not always strictly selective)[ref]. Besides *hCA II*, the transmembrane *hCA IX* and *hCA XII* isoforms are recognized markers in several types of cancer [REF]. In particular, *hCA IX* was found to be specifically overexpressed in hypoxic tumors and has recently raised particular interest as an anticancer drug target [REF][REF]. Moreover, *hCA XII* is up-regulated in many human cancers, including renal, pancreatic, brain, ovarian and lung cancers [ref][ref].

Estrogen Receptors are established anticancer drug targets. Antagonists of these proteins such as tamoxifen still represent the treatment of choice for estrogen responsive (ER α /ER β +) breast tumors [ref]. In particular, ER β plays a protective role by negatively modulating the ER α -driven uncontrolled proliferation in this type of malignancies, although ER β expression levels decrease during disease progression [ref]. However, ER β exerts a proliferative effect in ER α -negative tumors, in which it is often significantly over-expressed [refs]. Interestingly, expression levels of different *hCAs* and ERs are mutually linked in several types of cancer [REFS]. In particular, a direct correlation between the expression of ER α and *hCA XII* has been observed in ER α -positive breast cancer [REFS]. Moreover, the *hCA IX*-mediated acidosis of the tumor micro-environment has been associated with reduced efficacy of anti-estrogen therapy [ref]. In addition, *hCA IX* plays a major role in enhancing the aggressive phenotype of ER α -negative tumors [ref]. Based on these evidences, we reasoned that the combined modulation of selected *hCA* and ER proteins with a single multi-target ligand could provide improved anticancer activity with respect to single-target compounds. In particular, a molecule endowed with *hCA IX*/ER α or *hCA XII*/ER α dual modulatory activity would be useful for the treatment of estrogen responsive tumors, while a *hCA IX*/ER β dual antagonist might represent a valuable therapeutic agent in ER α -negative tumors such as Triple Negative Breast Cancer (TNBC).

Building on this rationale, we performed extensive structure-based analyses by docking the more promising hexahydrocyclopenta[*c*]quinoline compounds resulting from the ligand-based calculations into representative crystal structures of *hCAs* and ERs. Molecular dynamics simulations of the resulting complexes were also performed. Starting from our initial hit (\pm)-**3a**

(MOD23), we synthesized two additional derivatives, i.e. compounds (\pm)-**3b** (MOD66) and (\pm)-**3c** (MOD67) that contained both the hCA and ER pharmacophores at the appropriate location. Interestingly, the two compounds proved to modulate both cancer-related hCA IX and hCA XII carbonic anhydrase isoforms and ER activity. Crystal structures of key compounds bound to hCA II and ER α were then solved in order to provide structural insights into the observed activities and suggest possible improvements.

Obtaining a dual modulator of two very different targets as hCA and ERs is a challenge, especially when low molecular weight ligands based on integrated dual pharmacophores are looked for. This work moves the first step toward demonstrating that obtaining dual modulators of hCA and ER is feasible, and provides the structural grounds for their hit to lead optimization.

RESULTS AND DISCUSSION

Ligand-based virtual screening on the DrugBank, PDB and ChEMBL databases

The similarity profile of the investigated hexahydrocyclopenta[*c*]quinoline compounds was evaluated with respect to ligands reported in the DrugBank (approved or investigational drugs), PDB (ligands with 3D structural information) and ChEMBL (molecules with known activity annotations) databases, as described in the methods section. The 2D and 3D similarity analyses allowed the identification of a set of potential biological targets for the compounds of the investigated library. The 10 top-ranking targets resulting from comparison of the ligand-based similarity profiles and reported bioactivity annotations are summarized in **Table 1**. The complete list of targets is reported in **Table S1** of the **Supporting Information**.

Table: Top 10 targets resulting from the 3D ligand-based analyses on the DrugBank, PDB and ChEMBL databases. For each target, the number of hexahydrocyclopenta[*c*]quinoline molecules (queries) and of database entries identified by the ligand-based analyses on DrugBank, PDB and ChEMBL are reported (*TanimotoCombo* ≥ 1.4). Targets are sorted according to the number of identified active versus inactive ChEMBL compounds (activity threshold = 10 μ M). Targets that emerged from all the three databases are highlighted in bold.

UniProt ID	Target name	Query molecules	DrugBank entries	PDB entries	ChEMBL entries (inactives)
P00519	Tyrosine-protein kinase ABL1	2	0	0	34 (1)
P00915	Carbonic anhydrase 1	14	0	0	17 (0)
P52732	Kinesin-like protein KIF11	21	0	0	28 (6)
P00918	Carbonic anhydrase II	14	1	1	18 (1)
Q16790	Carbonic anhydrase IX	1	0	0	9 (0)
P30542	Adenosine receptor A1	1	0	0	6 (0)
P03372	Estrogen receptor α	22	1	1	7 (1)
O43570	Carbonic anhydrase XII	2	0	0	7 (1)
Q92731	Estrogen receptor β	4	1	1	5 (0)
P29274	Adenosine receptor A2a	1	0	0	7 (1)

According to the ligand-based results, *human* Carbonic Anhydrases (hCA) II, IX, XII and Estrogen Receptor (ER) subtypes α and β emerged as potential targets of some of the investigated compounds. In particular, a significant degree of similarity was identified between compound (\pm)-**3a** (MOD_23) and the DB07476 DrugBank entry [ref]. The latter compound is also reported in the PDB (in complex with hCA II, PDB accession code: 2HD6) [ref], and in ChEMBL (ChEMBL ID: ChEMBL386049) as a nanomolar inhibitor of several hCAs (e.g., isoforms I, II, IX and XII) [ref]. Moreover, compound MOD_63 resulted highly similar to the DrugBank compound DB07933 [ref], which corresponds to a low nanomolar modulator of the Estrogen Receptors. The latter compound is also reported in the PDB (PDB accession codes: 2I0J and 2I0G, in complex with ER α and ER β , respectively) [ref] and in ChEMBL (ChEMBL ID: ChEMBL278703) [ref]. **Figure 1** shows the ligand-based alignments predicted for **3a** (MOD_23) and MOD_63 with DB07476 and DB07933, respectively.

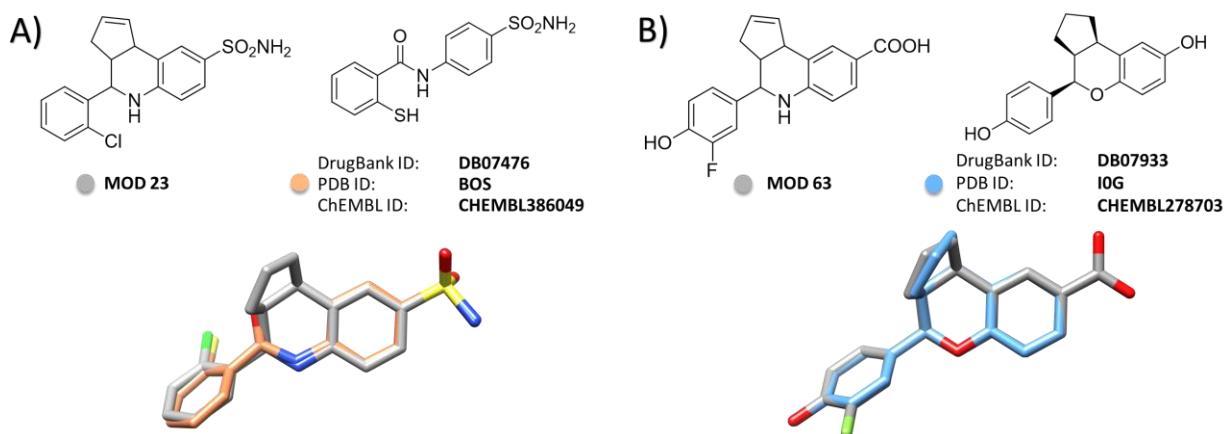


Figure 1. Ligand-based alignments predicted with ROCS. 2D chemical structures (top) and ligand-based alignments (bottom) are reported for A) DB07476 and MOD_23; B) DB07933 and MOD_63. The DB07476 and DB07933 ligands are represented as orange and blue sticks, respectively. The investigated MOD_23 and MOD_63 compounds are represented as grey sticks.

Considering that hCAs and ERs are known to contribute to tumor cell growth and development of drug resistance in several types of cancer, their expression levels being mutually regulated [ref], the hexahydrocyclopenta[*c*]quinoline scaffold may represent a valuable starting point for the development of dual ligands with potentially improved therapeutic activity.

Design of hCA/ER dual modulators

The design of hCA/ER dual modulators based on the hexahydrocyclopenta[*c*]quinoline scaffold was guided by ligand-based alignments predicted for **3a** MOD_23 and MOD_63 and by docking into representative crystal structures of both targets (**Figure 2**). In particular, MOD_23 bears a sulfonamide moiety able to coordinate the catalytic zinc ion of hCAs [ref], and MOD_63 possesses an hydroxyl group crucial for binding to ERs [REF][REF].

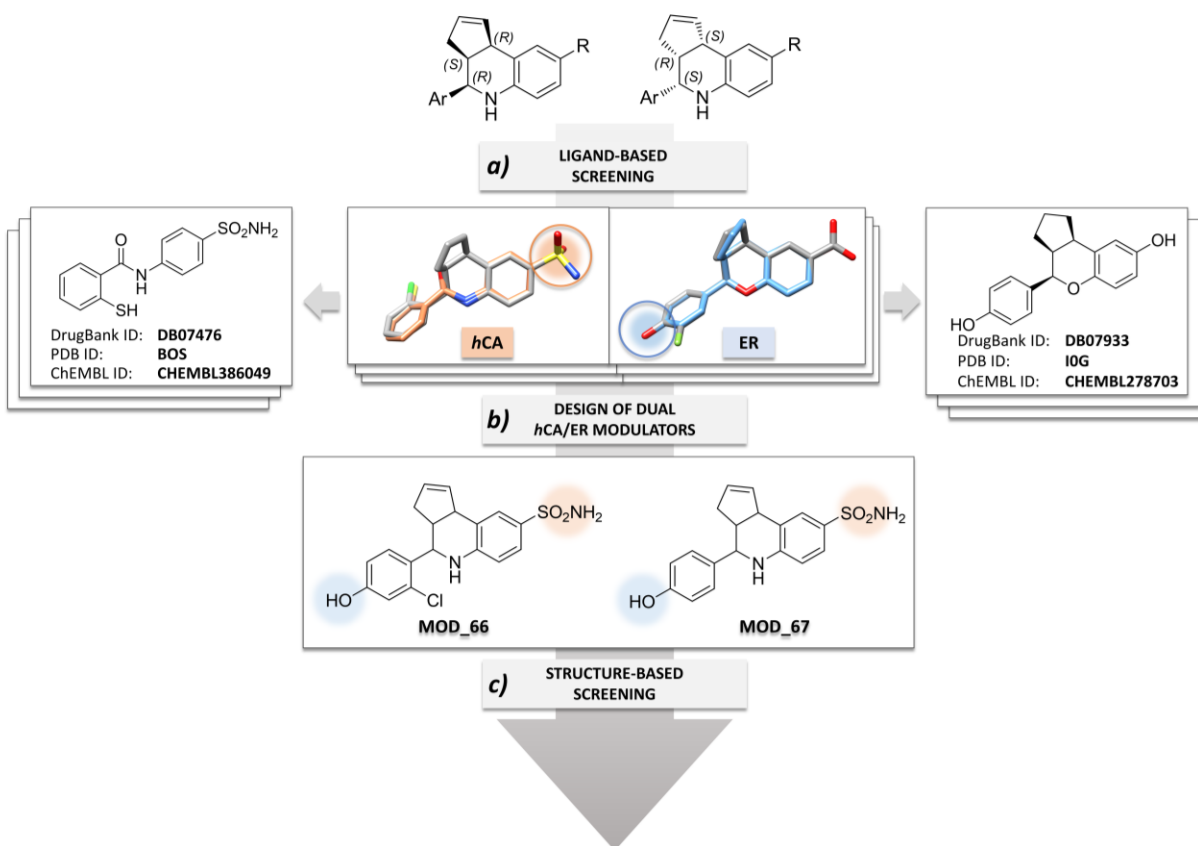


Figure 2: Workflow adopted for the design of dual hCA/ER modulators. From top to bottom: (a) library molecules were subjected to a ligand-based screening. Modulators of hCA and ER emerged as highly similar to several of the investigated compounds; significant alignments are displayed as sticks (orange sticks: DB0746 entry; light blue sticks: DB07933 entry; gray sticks: top-scoring investigated molecules). (b) Two new derivatives were designed with the aim to obtain dual modulators of both hCA and ER. Key pharmacophoric moieties are highlighted with orange and blue circles for hCA and ER, respectively. (c) A structure-based screening was performed on the two newly designed molecules endowed with potential dual hCA/ER modulation activity.

Therefore, two novel hexahydrocyclopenta[*c*]quinoline derivatives ((±)-**3b** **MOD_66** and (±)-**3c** **MOD_67** in **Figure 2**) were designed by incorporating both the sulfonamide moiety (hCA) and the *p*-hydroxyl group (ER) at the appropriate locations. Then, docking calculations were performed to evaluate the ligand-protein complementarity with the selected hCAs (isoforms I, II, IX and XII) [ref] and ER (α and β subtypes) crystal structures (see **Methods**, and **Table S2** in the **Supporting Information**).

Visual inspection of the resulting docking poses (**Figure S2** in the **Supporting Information**) showed that the binding mode of **3b** **MOD_66** and **3c** **MOD_67** is consistent with those of previously reported sulfonamide hCA inhibitors [ref], the sulfonamide moiety coordinating the catalytic zinc ion. The hydroxyl group of the phenyl ring established hydrogen bonds with the side

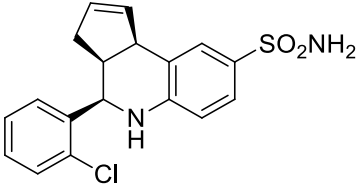
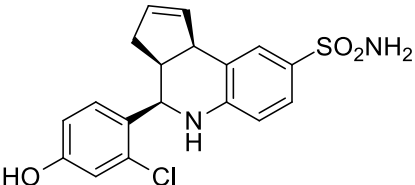
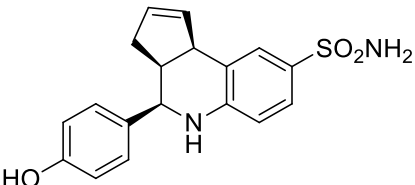
chains of the Asp131 (*hCA IX*) or Ser130 (*hCA XII*) residues (**Figure S2** of the **Supporting Information**). Such hydrogen bonds are not possible in *hCA I* and *hCA II*, the corresponding residues being Ala131 and Gly132, respectively. Moreover, docking calculations revealed steric clashes between the cyclopentene moiety and the Leu131 (*hCA I*) or Phe131 (*hCA II*), the latter residues being bulkier than the corresponding Val130 (*hCA IX*) and Ala129 (*hCA XII*) residues. Altogether, these differences indicate that the newly designed compounds may display selectivity for the cancer-related isoforms *hCA IX* and *hCA XII*.

As for the ER α and ER β docking complexes (**Figure S3** in the **Supporting Information**), the *p*-hydroxyl group of **MOD_66** and **MOD_67** is engaged in a hydrogen bond network with Glu353 and Arg394 in ER α , and with the corresponding Glu305 and Arg346 in ER β , thus mimicking the A-ring phenol of estradiol and the phenol ring of 4-hydroxytamoxifen. Moreover, the sulfonamide moiety was predicted to establish hydrogen bonds with the His524 and His475 residues in ER α and ER β , respectively. Both compounds were predicted to accommodate near the Met343, Leu346, Ala350, Leu384, Leu387, Leu391 and Phe404 residues in ER α (PDB code 2AYR), and Met295, Leu298, Ala302, Met336, Leu339, Leu343 and Phe356 in ER β (PDB code 1L2J), establishing hydrophobic contacts similar to those observed for other crystallographic ligands [ref].

Biological evaluation of compounds (\pm)-3a-3c** **MOD_23/66/67** on carbonic anhydrase isoforms**

Compounds (\pm)-**3a-3c** were tested for their *in vitro* inhibition of *hCA* isoforms I, II, IX and XII, using acetazolamide (AAZ) as a positive control. Interestingly, all the investigated compounds potentially inhibited *hCAs*, albeit with different selectivity profiles (**Table 2**). In particular, while (\pm)-**3a** **MOD_23** displayed selectivity for *hCA II* and to a lesser extent *hCA I*, (\pm)-**3c** **MOD_67** was selective for *hCA I* and to a lesser extent *hCA II*. Both compounds were relatively less active against the cancer-related isoforms IX and XII. On the contrary, (\pm)-**3b** **MOD_66** turned out to be a potent inhibitor of *hCA IX* and *hCA XII*, with 10-fold selectivity with respect to *hCA I* and *hCA II*.

Table 2: Inhibitory activities (K_i , nM) of the synthesized hexahydrocyclopenta[*c*]quinoline compounds on different hCA isoforms.

Cmp		K_i (nM)*			
		<i>hCA I</i>	<i>hCA II</i>	<i>hCA IX</i>	<i>hCA XII</i>
(±)- 3a		73.6	2.2	2089	255.8
MOD_23					
(±)- 3b		706.2	539.1	56.3	78.8
MOD_66					
(±)- 3c		6.7	94.6	803.5	346.2
MOD_67					
AAZ		250	12.1	25.8	5.7

* Mean from 3 different assays, by a stopped flow technique (errors were in the range of $\pm 5-10$ % of the reported values).

Crystal structures of compounds (±)-**3a-3c** in complex with hCA II

To experimentally evaluate the binding mode of (±)-**3a-3c** **MOD_23**, **MOD_66** and **MOD_67** in hCA, crystallographic experiments were performed on hCA II (**Table S3** of **Supporting Information**).

The analysis of the initial Fo–Fc electron density map of the active site was fully compatible with the presence of the inhibitors. The predicted stereochemistry of the crystallized compounds was consistent with the 3*aR*, 4*S*, 9*bS* ~~12*S*, 16*R*, 17*S*~~ isomer. While the sulfonamide-quinoline moiety displayed a well defined electron density, the cyclopentene moieties and the tail aromatic rings bearing a chlorine and/or hydroxyl substituent appeared to be not fully resolved (**Figure 3A-C**).

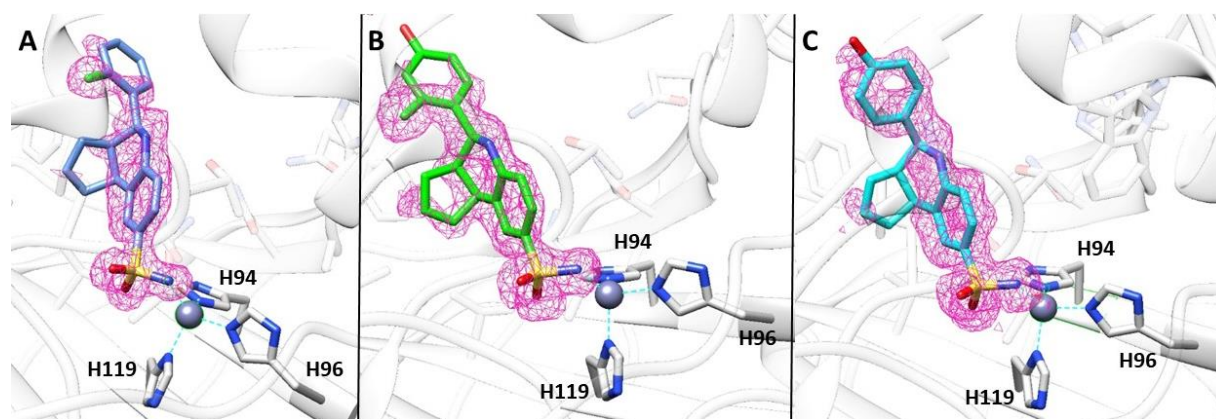
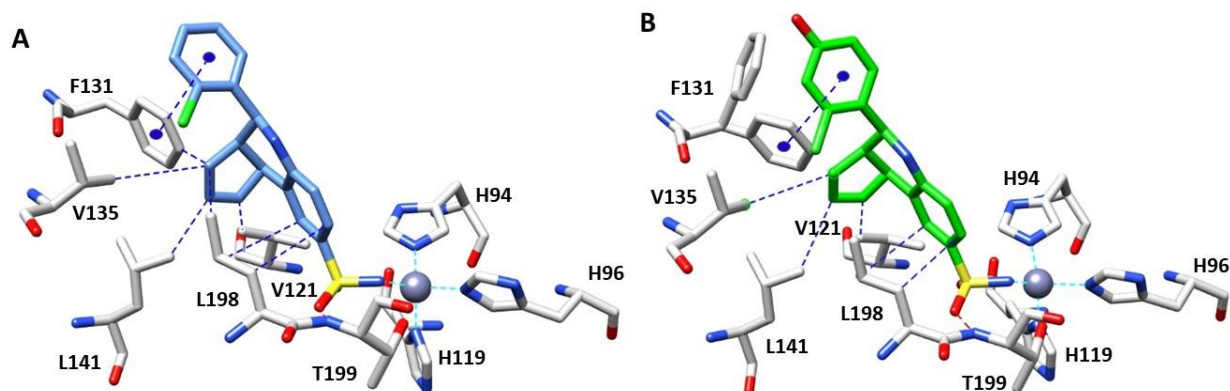


Figure 3: Active site region of *hCA II*/**MOD_23** (PDB: 6SX9) (A), *hCA II*/**MOD_66** (PDB: 6SYB) (B) and *hCA II*/**MOD_67** (PDB: 6SYS) (C) adducts. Inhibitors showed as σA -weighted $|Fo-Fc|$ density map at 2.0σ . Zinc coordination is also shown.

In each protein–ligand complex, the sulfonamide moiety directly binds the zinc ion in the active site with a tetrahedral geometry involving His94, His96 and His119, as already observed for other inhibitors [REF]. An additional hydrogen bond interaction, often observed for compounds with this ZBG, is engaged between the sulfonamide oxygen and the nitrogen of residue Thr199, which helps stabilizing the complexes [REF]. As shown in **Figure 4**, the quinoline-sulfonamide scaffold of **MOD_23**, **MOD_66** and **MOD_67** established favorable van der Waals interactions with the side chain of residue Leu198. In addition, several hydrophobic interactions are engaged by the cyclopentene moiety of the inhibitors with the Val135, Leu141 and Val121 residues of *hCA II*. The *o*-Cl substituent on the aromatic ring plays a fundamental role for the ligand-protein interactions. In fact, its presence on the aromatic ring moves the tail in a favorable position to engage a π -stacking (**Figure 4**, panels A and B). On the other hand, the aromatic ring of **MOD_67** adopts a different orientation and makes van der Waals interactions with the side chains of residues Pro202, Leu198 and Phe131 (**Figure 4**, panels C and D). In addition, the quinoline-sulfonamide scaffold of **MOD_67** shows a water bridge interaction between the quinoline nitrogen and that of the side chain of Gln92.



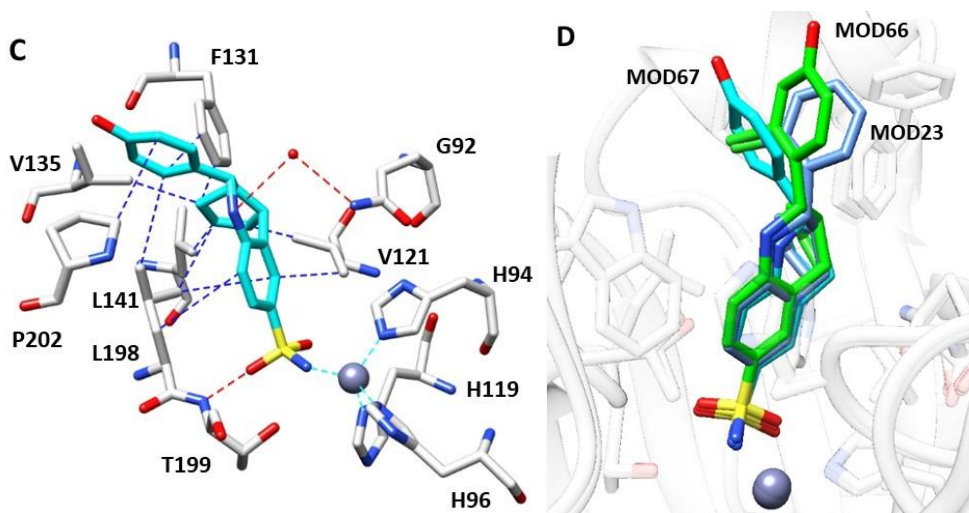


Figure 4: (A) Active site region of hCA II/**MOD_23**. (B) Active site region of hCA II/**MOD_66**. (C) Active site region of hCA II/**MOD_67** adducts. Hydrogen bonds and van der Waals interactions are labelled and shown in red and blue, respectively. (D) Structural superposition between **MOD_23** (cornflower blue), **MOD_66** (green) and **MOD_67** (cyan) bound to the hCA II active site.

Overall, the hexahydrocyclopenta[*c*]quinoline scaffold proved to be able to bind the hCA II active site. Moreover, the different substituents on the aromatic ring of the investigated inhibitors demonstrated to play a role in determining the potency and selectivity against different hCA isoforms, which is an important aspect to take into account for the design of analogues with optimized dual hCA/ER activity.

Anti-cancer activity of the investigated compounds

We evaluated the anti-cancer properties of compounds (\pm)-**3a-3c** **MOD_23**, **MOD_66**, and **MOD_67** using cell culture models. In about 70% of breast cancers cases, ER α drives tumor growth in response to estrogens and other environmental cues. In MCF7 cells (ER α +, ER β -), a model for estrogen-dependent breast cancer, all three compounds demonstrated anti-proliferative effects at micro-molar concentrations (**Figure 5A**). The natural estrogen, 17 β -estradiol (E2) completely reverses the growth-inhibitory effects of classical SERMs such as raloxifene (RAL). However, E2 partially reversed the effect of the most potent compound (\pm)-**3c** **MOD_67**, as well as that of (\pm)-**3a** **MOD_23**, but not (\pm)-**3b** **MOD_66** (**Figure 5B**), suggesting that ER α mediates some of their anti-proliferative effects on breast cancer cells. Consistent with this, the anti-proliferative effects of all three compounds were considerably reduced in the triple-negative breast cancer (TNBC) model, MDA-MB231 cells (ER α -, ER β +) compared to MCF7 cells (**Figure 5A**).

In prostate cancer, ER α facilitates disease progression (PMID: [25415230](#)), while ER β is a putative tumor suppressor. However, another steroid receptor, the androgen receptor (AR), predominantly drives cell proliferation in the early stages of prostate cancer progression, and a naturally occurring AR-T877A ligand-binding domain mutation found in prostate cancer cells, allows AR to bind and respond to estrogens.

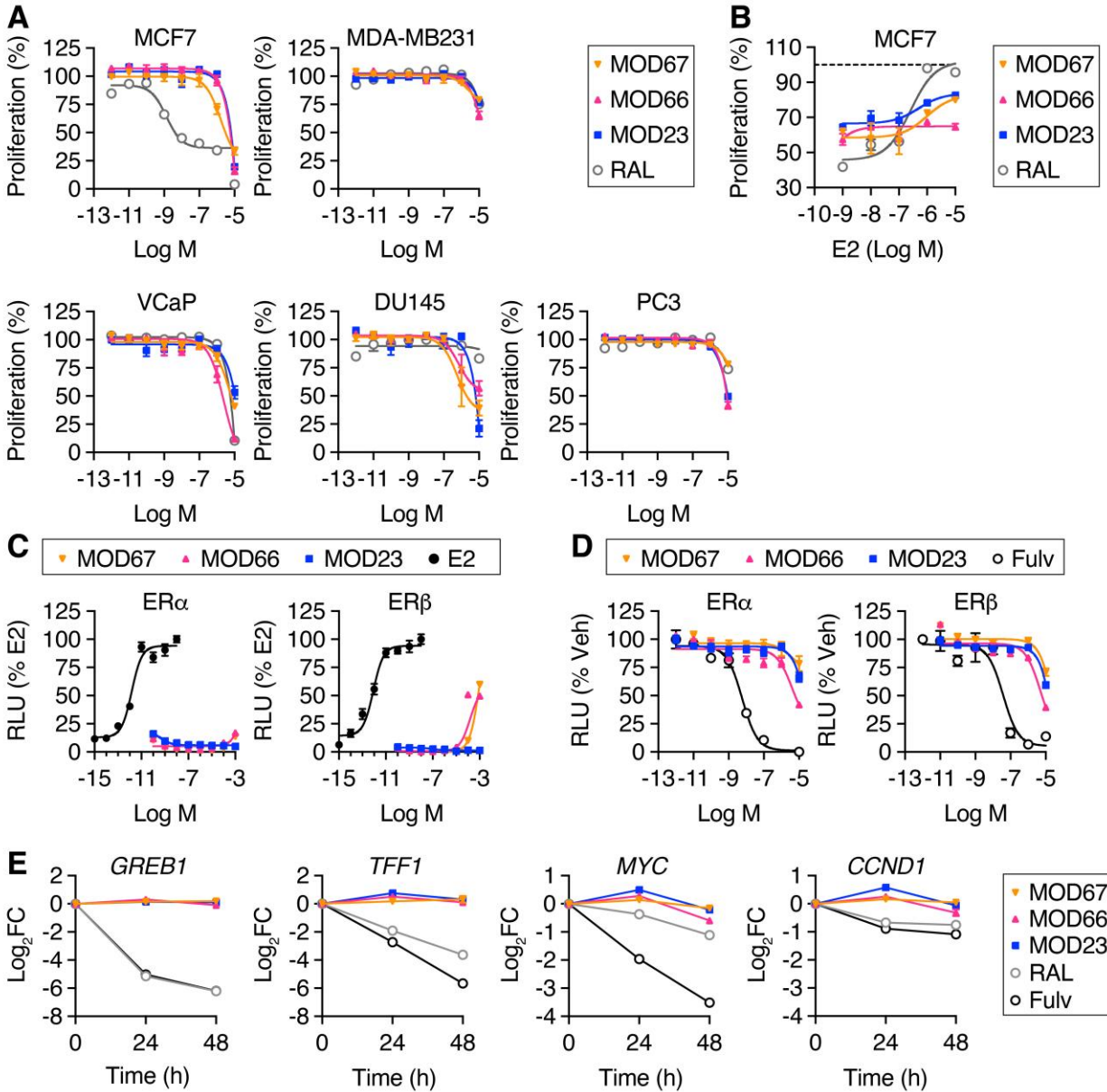


Figure 5. Evaluation of biological activity in cancer cells. (A) Effects on cell proliferation. Different breast and prostate cancer cell types were treated with the indicated compounds for 5 days. The amount of cell proliferation is shown as a percentage of DMSO (vehicle)-treated cells \pm SEM ($n=6$). (B) Blockade of ER-mediated effects. MCF7 cells were stimulated with 5 μ M MOD_23, MOD_66, or MOD_67, or 1 μ M raloxifene, and increasing doses of E2. The amount of cell proliferation is shown as a percentage of vehicle-treated cells \pm SEM ($n=4$). (C–D) Modulation of ER activity. Steroid-deprived 293T cells transfected with a 3xERE-luciferase reporter and an ER α or ER β expression plasmid were treated with the indicated compounds C) alone, or D) in combination with 1 nM E2. Luciferase activity is shown as a percentage of E2-treated cells \pm SEM ($n=6$). (E) ER-target gene expression. MCF7 cells were treated with vehicle, 5 μ M MOD_23, MOD_66, or MOD_67, 1 μ M raloxifene, or 0.1 μ M fulvestrant. GREB1, TFF1, MYC, and CCND1 mRNA levels were compared by qPCR and normalized to AKT1. Datapoints represent log-transformed fold changes relative to vehicle-treated cells ($\text{Log}_2\text{FC} \pm \text{SEM}$; $n=3$).

Here, we examined proliferative effects on VCaP cells, an androgen-sensitive prostate cancer model with wild type AR (PMID: 16520280), and on DU145 and PC3 cells, both of which are androgen-insensitive prostate cancer models without AR (PMID: 32132580), to exclude potential AR-mediated proliferative effects. Micro-molar doses of (±)-**3a-3c** had anti-proliferative effects on VCaP cells where (±)-**3b** MOD_66 was the most potent, on DU145 cells where (±)-**3c** MOD_67 was the most potent, and on PC3 cells where (±)-**3c** MOD_67 was the least potent (Figure 5A). These results suggest that the anti-proliferative effects of (±)-**3a-3c** MOD_23/66/67 on cancer cells are achieved at micro-molar concentrations that are potentially high enough to modulate multiple factors, including carbonic anhydrase and ER.

Biological Evaluation of compounds (±)-**3a-3c** on Estrogen Receptor Activity

To compare effects on ER activity, we performed a 3xERE-driven luciferase transactivation assay in agonist mode using HEK-293T cells. In this assay, none of the compounds activated ER α , but (±)-**3b** MOD_66 and (±)-**3c** MOD_67 partially activated ER β compared to the natural estrogen, 17 β -estradiol (E2). (±)-**3b** MOD_66 (EC₅₀: 0.15 mM) was a more potent partial ER β agonist than (±)-**3c** MOD_67 (EC₅₀: ~1.15 mM), while (±)-**3a** MOD_23 had no detectable ER β agonist activity (Figure 5C).

Effects on estradiol-induced ER α and ER β activities were also compared by 3xERE-driven luciferase assay in antagonist mode using HEK-293T cells. Unlike the classical SERD, fulvestrant (Fulv), which completely abolished the agonist activity of E2, (±)-**3b** MOD_66 reduced ER α - and ER β -mediated activity of E2 by at least 58% and 60%, respectively (Figure 5D). (±)-**3a** MOD_23 was less potent than (±)-**3b** MOD_66, reducing ER α - and ER β -mediated activity of E2 by at least 34% and 40%, respectively (Figure 5D). (±)-**3c** MOD_67 was the least potent, reducing ER α - and ER β -mediated E2 activity by at least ~22% and 29%, respectively (Figure 5D). These results demonstrate that compounds (±)-**3a-3c** MOD_23/66/67 exhibit partial ER agonist/antagonist activities.

To further understand the anti-proliferative effects of these compounds, we examined their effects on the endogenous levels of estrogen-induced mRNAs in MCF7 cells. Raloxifene and fulvestrant strongly suppressed expression of canonical ER α -target genes such as *GREB1* and *TFF1/pS2*, and subtly downregulated non-canonical ER α -target genes such as the *MYC* oncogene, and *CCND1/Cyclin D1* (Figure 5E). In contrast, (±)-**3a-3c** MOD_23/66/67 did not have obvious inhibitory effects on expression of these genes (Figure 5E), indicating that their anti-proliferative effects on MCF7 cells are not due to antagonism of the classical ER α signaling pathway, but are probably achieved through an alternative mechanism.

Structural Basis for the Modulation of ER α Activity

To determine how **3b** MOD_66 and **3c** MOD_67 modulate ER activity, the X-ray crystal structures of the ER α -Y537S mutant ligand-binding domain (LBD) were solved in complex with **3a/3b** MOD_66/67 and a coactivator peptide (Table S3). Trials of **3a/3b** MOD_66/67 plus the ER α -

L372S/L536S double-mutant LBD that favors the 4-hydroxytamoxifen-bound LBD conformation failed to yield crystals. In each **3a/3b** MOD_66/67-bound LBD subunit, the switch helix, h12, docks across helices h3 and h11, while a coactivator peptide of NCOA2/SRC-2 binds the AF-2 surface (**Figure 6**, panel A). Thus, **3b** MOD_66 and **3c** MOD_67 stabilized active ER α LBD conformations, indicative of ER agonists.

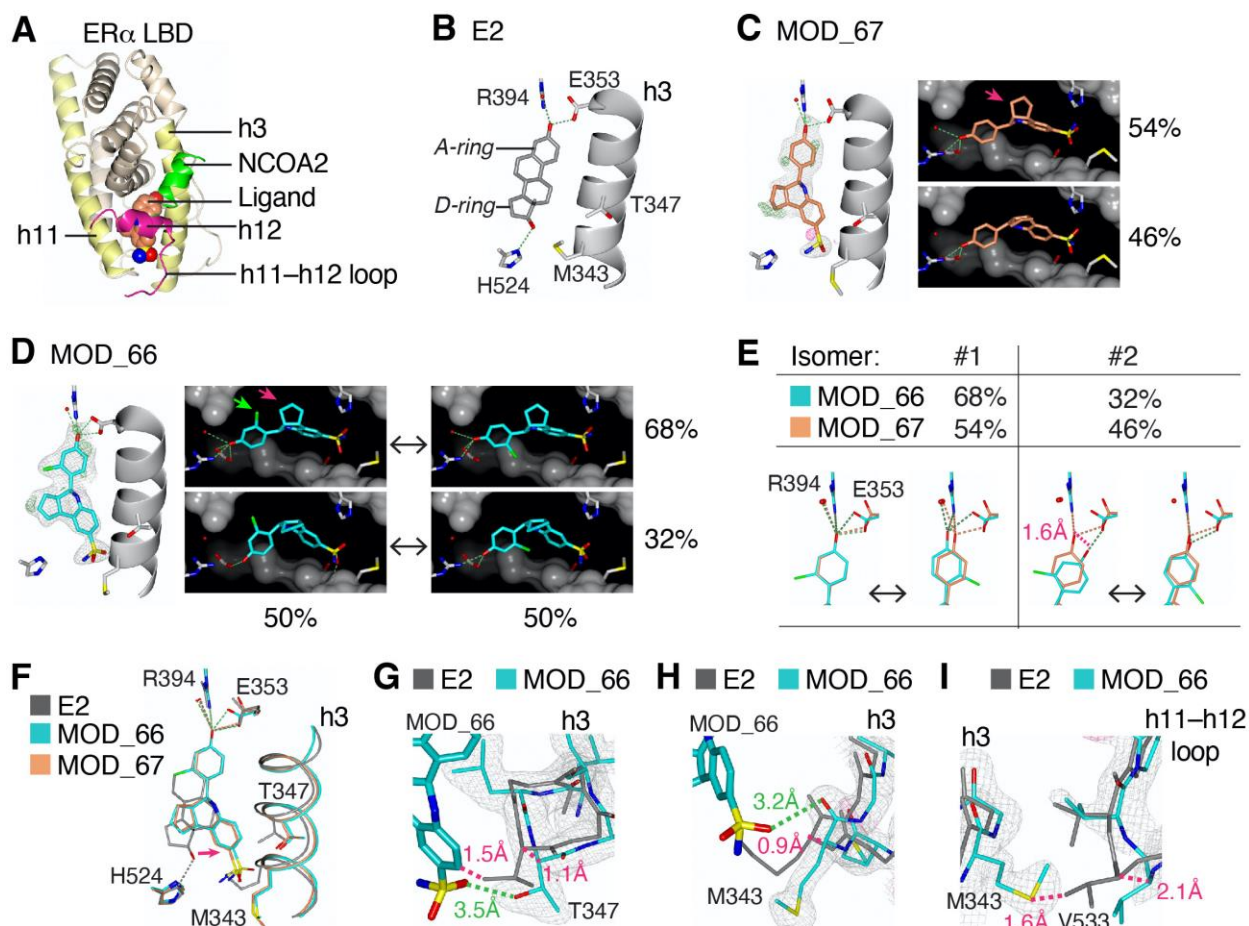


Figure 6. Structural basis for the modulation of ER α activity by MOD_66 and MOD_67. (A) Overview of the crystal structure of MOD_66/67-bound ER α LBD subunit; (B–D) Crystal structures of the active ER α LBD showing the binding modes of (B) E2 (3UUD.pdb), (C) MOD_67 stereoisomers, and (D) alternate conformers of MOD_66 stereoisomers; (E–F) Superposed LBDs showing (E) isomer-selective rotation of the MOD_66 A-ring, and (F) differences in D-ring positioning (arrow) and h3 conformation; (G–H) Details of MOD_66-induced conformational changes at h3 (G) Thr347 and (H) Met343; (I) The MOD_66-induced allosteric signal is transmitted from h3 to the h11–h12 loop. $2F_o - F_c$ electron density and $F_o - F_c$ difference maps were contoured at 1σ and 3σ , respectively.

Consistent with its predicted binding mode (**Figure S3** in the **Supporting Information**), the phenol ring of **3c** MOD_67 mimics the A-ring of estradiol (PDB accession code: 3UUD, PMID: 22927406) and forms hydrogen bonds with Glu353 and Arg394. Estradiol forms an additional hydrogen bond with h11 at His524 (**Figure 6**, panel B), but the sulfonamide moiety of **3c** MOD_67

targets a distinct h3 epitope near Thr347 and Met343 (**Figure 6**, panel C). In the LBD subunit with a clearer ligand electron density, both **3c MOD_67** stereoisomers occupied the ligand-binding pocket in this orientation, and with similar occupancies, suggesting that active ER α LBD forms equally stable complexes with **3c MOD_67** isomers.

The **3b MOD_66**-bound LBD structure also contained two distinct stereoisomers, each of which adopts two distinct conformations inside the ligand-binding pocket (**Figure 6**, panel D). The A-ring hydroxyl group of each isomer forms hydrogen bonds with Glu353 and Arg394, while the chlorine atom is accommodated on either side of the pocket, 50% of the time (**Figure 6**, panel D). However, one **3b MOD_66** isomer occupied the pocket about 70% of the time, suggesting that the active ER α LBD forms a more stable complex with this isomer.

To understand this isomer-selectivity, we superposed the **3b MOD_66**- and **3c MOD_67**-bound LBDs and compared the A-ring interactions in each structure. We found that the pendant chlorine atom of the preferred **3b MOD_66** isomer (i.e. isomer #1) is accommodated with little-to-no rotation of the A-ring (**Figure 6**, panel E). In contrast, accommodation of the chlorine atom in isomer #2 required considerable rotation of A-ring, which moves the hydroxyl group ~ 1.6 Å away from the expected position, and destabilizes the H-bonding network, 50% of the time (**Figure 6**, panel E). **This isomer-selective rotation of the A-ring likely reduces the binding affinity of 3b MOD_66 isomer #2, and thereby favor the formation of ER α LBD complexes with 3b MOD66_isomer #1.**

To understand how **3b** and **3c MOD_66/67** affects ER α LBD structure, the estradiol-bound LBD, (PDB accession code: 3UUD, PMID: 22927406), was superposed on the **3b/3c MOD_66/67**-bound structures. Unlike the D-ring of E2 which is accommodated closer to h11 and stabilized by H-bonding with His524 (**Figure 6**, panel B), the D-rings of **3b/3c MOD-66/67** are accommodated near h3, where it is not stabilized by a H-bond, but distorts the LBD conformation (**Figure 6**, panel F). **To this end**, the D-ring and the attached sulfonamide moiety pushes h3 Thr347 and Met343 away from the pocket by ~ 1 Å (**Figure 6**, panels G–H). This forces Met343 to displace the h11–h12 loop residue, Val533 by ~ 2 Å from its estradiol-bound position (**Figure 6**, panel I). Together, these results suggest that **3b MOD_66** and **3c MOD_67** are partial ER agonists that bind inside the ligand-binding pocket, and directly dislocate h3 of the LBD to modulate receptor structure and activity by forming A-ring H-bonds with Glu353 and Arg394, but the lack of additional H-bonds and the clashes between the sulfonamide moiety and h3 residues, drastically reduces their ER α binding affinity.

The h3-mediated perturbation of the h11–h12 loop by **3b/3c MOD_66/67** may affect h12 dynamics in solution and thereby alter coactivator binding at the AF-2 surface. However, their predominant binding mode where they wrap around h3, forming hydrogen bonds with Glu353, Thr347, and/or the Met343 peptide link, is consistent with an alternative, *GREB1*-independent mechanism of action. Many other ER ligands that wrap around h3, including furan and 3,4-DTP analogs, modulate ER α -mediated cell proliferation through *GREB1*-independent pathways (PMID: 27107013). In addition, each **3b MOD_66** isomer binds the ER α LBD in more than one orientation, indicating that they exhibit ligand dynamics, which rewires the ER α signaling network and enables

ligand-dependent control of ER α -mediated cell proliferation through *GREB1*-independent pathways (PMIDs: 23524984, 27107013).

CONCLUSIONS

One of the challenges of single agent polypharmacology in rational drug design is that it requires optimization of the key pharmacophores for unrelated protein targets (PMID: 26283966, 29126724). The compounds presented here are hexahydrocyclopenta[*c*]quinoline analogs designed to inhibit carbonic anhydrase isoforms and modulate ER activity. These compounds bind inside the ER ligand-binding pocket using their A-ring pharmacophore to form H-bonds with the receptor. The key pharmacophore for the inhibition of carbonic anhydrase – the sulfonamide moiety at the other end of these compounds – also causes severe steric clashes that distort ER LBD structure and thereby modulates ER activity. These compounds are not potent ER modulators and exhibit multiple binding orientations in the LBD, consistent with poor ER binding affinities. Thus, further ligand optimization is required to develop analogs with more potent ER antagonist activities, without compromising the inhibition of carbonic anhydrase isoforms, which is high.

EXPERIMENTAL SECTION

Ligand-based virtual screenings on the DrugBank, PDB and ChEMBL databases

The investigated hexahydrocyclopenta[*c*]quinoline-based compounds [ref] were prepared by using the *LigPrep* (version 4.7.012) utility available within the Schrödinger Suite 2018-3 [ref]. In particular, all combinations of ionization states and tautomers potentially present at physiological pH were generated. Then, five conformers were generated for each of the pre-treated compounds by using the OMEGA software (version 3.0.1.2), with default settings [ref].

Ligands with already reported activity annotations were firstly retrieved from the DrugBank (www.drugbank.ca, accessed on November 13, 2018), PDB (www.rcsb.org, accessed on November 20, 2018) and ChEMBL (www.ebi.ac.uk/chembl/, accessed on November 13, 2018) databases [ref]. These databases provide structural, chemical and biological information related to ligands and targets, and are of high relevance for drug discovery and repurposing [ref].

The DrugBank and PDB databases were filtered by retaining only compounds with a molecular weight between 120 and 900 Da, and at least six heavy atoms and one aromatic ring. Compounds containing toxic non-metals and transition metals were discarded. The ChEMBL database was filtered to retain only compounds with reported activity annotations expressed as IC₅₀, EC₅₀, K_i or K_d, and resulting from binding experiments on single or homologous proteins.

Then, the filtered databases were pre-treated by using *LigPrep* (version 4.7.012) to generate all combinations of ionization states and tautomers potentially present at a pH of 7±2 [ref]. Moreover, all stereoisomers were generated for compounds with undefined chirality. Afterwards, up to 400 conformers were generated for each ligand in the prepared databases by using the OMEGA software (version 3.0.1.2) with default settings [ref].

3D ligand-based virtual screenings were performed by using the ROCS software (version 3.2.2.2) to assess the similarity profile of each of the investigated hexahydrocyclopenta[*c*]quinoline-based

compounds against the prepared DrugBank, PDB and ChEMBL databases. Results of the 3D ligand-based virtual screenings were firstly filtered to retain similarity records with a TanimotoCombo (*i.e.*, the sum of the ShapeTanimoto and ColorTanimoto coefficients) score higher than 1.4, which is considered an appropriate similarity threshold to classify ligands potentially exhibiting comparable bioactivity [ref]. Then, filtered records were matched with structural and bioactivity annotations reported in the DrugBank, PDB and ChEMBL databases. The identified targets were prioritized according to the number of known active *versus* inactive ligands that were similar to the investigated hexahydrocyclopenta[*c*]quinoline compounds. An activity threshold of 10 μM was used to discriminate between active and inactive compounds.

Structure-based calculations on selected targets

Docking calculations were performed on multiple conformations of the top scoring targets identified by the ligand-based analyses described above [REF]. The selection of suitable protein conformations was guided by the previously evaluated ligand similarities (see **Table S2** in the **Supporting Information**). In fact, the more similar the screened compounds are to co-crystallized ligands that bind to a given protein conformation, the more accurate docking calculations are in discriminating active from inactive compounds [ref]. X-ray crystal structures of *hCA* and ER proteins, whose ligands showed a high degree of similarity with at least one of the investigated hexahydrocyclopenta[*c*]quinoline-based compounds, were downloaded from the PDB (**Table S2** of the **Supporting Information**) and pre-processed with the *Protein Preparation Wizard* utility of Schrödinger Suite 2018-3 [ref]. In particular, atom types and bond connectivity were fixed, and missing residue side chains were rebuilt by using *Prime* (version 5.3) [ref]. Then, hydrogen atoms were added to the pre-treated protein structures and their coordinates were optimized. Afterwards, ions, solvent and water molecules with less than 3 interactions with the ligand and/or protein were removed. The zinc ion required for catalytic activity in the carbonic anhydrases [ref] was retained, while other metals were removed. Docking calculations were performed on the active site of selected protein conformations by using the Standard Precision (SP) mode of the *Glide* software (version 8.0.012) [ref]. Receptor grids were centered on the centroids of the bound ligands with default settings (outer box of 10 Å \times 10 Å \times 10 Å). Docking models were validated by redocking the co-crystallized ligands into their parent crystal structures. All redocking experiments correctly reproduced the crystallographic poses with a RMSD lower than 2 Å.

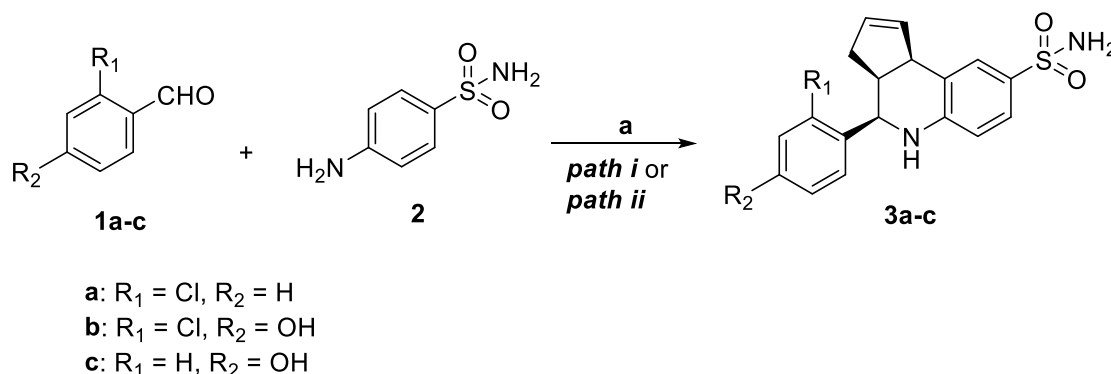
The hexahydrocyclopenta[*c*]quinoline compounds were prepared with *LigPrep* for the docking calculations [ref]. In particular, all combinations of ionization states and tautomers potentially present at physiological pH in aqueous solution (pH = 7 \pm 2) were firstly generated [ref]. As the *hCA* proteins contain a zinc ion in the binding site, additional metal binding states were also calculated. Then, the pre-treated structures were minimized according to the OPLS3e force field and docked into the validated receptor grids [ref]. Finally, the predicted poses were visually inspected and compared with those of already reported crystallographic complexes.

Flexible docking calculations were also performed by using the standard *Induced Fit Docking* (IFD) protocol implemented in the Schrödinger Suite 2018-3 [ref] to further account for potential

binding site flexibility. The predicted binding poses were visually inspected and compared with those predicted by *Glide* rigid docking, leading to a final selection of candidates for synthesis and biological testing.

Chemical synthesis of the newly designed derivatives

The synthesis of the racemic compounds (\pm)-**3a-c** (Scheme 1) is based on a multicomponent Povarov reaction previously described in the literature [REF]. In particular, AlCl_3 (Scheme 1, path ii) is essential to promote both the imine formation and the cycloaddition in the presence of the electron rich aldehydes **1b-c**. NMR analysis comparison of the newly synthesized racemic derivatives (\pm)-**3b** MOD_66 and (\pm)-**3c** MOD_67 with reported related compounds [REF], confirmed the relative configuration of (\pm)-**3 b-c**.



Scheme 1. Reagents and Conditions: a) path i-Synthesis of (\pm)-**3a**: cyclopentadiene, TFA, MeCN, 0 °C, 2h, 75%; path ii-Synthesis of (\pm)-**3a,b**: cyclopentadiene, AlCl_3 , MeCN, 0 °C, 8h, 51 – 64b%.

In vitro evaluation of the compounds on hCA isoforms

ADD PROTOCOL

In vitro evaluation against ER α and ER β

Cell culture

HEK-293T, MCF7, MDA-MB231, DU145, and VCaP cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) (ThermoFisher Scientific, Inc. Waltham, MA; cat no. 11995073), while PC3 cells were cultured in Roswell Park Memorial Institute (RPMI) 1640 medium. Both media were supplemented with 10% fetal bovine serum (FBS) (Sigma-Aldrich, Inc. St. Louis, MO; cat no. F0926), GlutaMAX™ (ThermoFisher Scientific; cat no. 35050061), MEM non-essential amino acids (ThermoFisher Scientific; cat no. 11140050), and an antibiotic mixture containing penicillin, streptomycin, and neomycin (ThermoFisher Scientific; cat no. 15640055). The cells were maintained at 37°C and 5% CO_2 .

Luciferase assay

HEK-293T cells in a 10 cm dish containing 10 ml of phenol red-free DMEM (ThermoFisher Scientific; cat no. 31053028) plus 10% charcoal-stripped FBS (ThermoFisher Scientific; cat no. A3382101) were co-transfected with 10 µg of 3xERE-driven luciferase reporter plasmid (3xERE-Luc) plus 5 µg of wildtype human ER α or ER β expression plasmids, using 45 µl of TransIT[®]-LT1 reagent (Mirus Bio LLC, Madison, WI; cat no. MIR 2300). The next day, the cells were transferred to either a 96-well plate and treated with test compounds, or a 384-well plate where the test compounds were added using a Biomek NX^P 100-nl pintool (Beckman Coulter, Inc. Brea, CA). Luciferase activity was measured the next day using the britelite plus reporter gene assay system (PerkinElmer, Waltham, MA; cat no. 6066761).

Cell proliferation assay

MCF7, MDA-MB231, DU145, and VCaP cells were suspended in phenol red-free DMEM plus 10% FBS. PC3 cells were suspended in phenol red-free RPMI 1640 medium plus 10% FBS. 1,000 cells were placed in each well of a 384-well plate. After about 4 h, the test compounds were added using a Biomek NX^P 100-nl pintool (Beckman Coulter). For 96-well assays, 40,000 cells/ml suspensions were treated with test compounds and seeded at a density of 4,000 cells/well. The relative amount of proliferation was determined 5 days later using the CellTiter-Glo[®] Luminescent cell viability assay (Promega, cat no. G7572).

Quantitative RT-PCR (qPCR)

MCF7 cells in a 6-well plate containing 2 ml/well of phenol-red free DMEM plus 10% whole FBS, were treated with test compounds. Total RNA was isolated using the RNeasy[®] mini kit (QIAGEN, Venlo, Netherlands; cat no. 74106) with on-column DNase digest. cDNA was synthesized in 20 µl reactions containing 1 µg of RNA, using the TaqMan[™] reverse transcription kit (ThermoFisher Scientific; cat no. 4304134), and then diluted in an equal volume of nuclease-free, molecular grade water. The samples were then analyzed by real-time PCR. Each reaction contained 1 µl of diluted cDNA, 5 µl of TaqMan[™] gene expression master mix (ThermoFisher Scientific; cat no. 4369016), 3.5 µl of nuclease-free water, and 0.5 µl of *GREB1* (Hs00536409_m1), *TFF1* (Hs00170216_m1), *MYC* (Hs00153408_m1), *CCND1* (Hs00277039_m1), or *AKT1* (Hs00178289_m1) TaqMan[™] gene expression assay primers and probes (ThermoFisher Scientific; cat no. 4331182). Relative mRNA levels were normalized to *AKT1* using the $\Delta\Delta C_T$ method, as previously described (PMID: 18265376).

Crystallization and X-ray data collection of *hCA* complexes

Crystals were obtained using the hanging drop vapor diffusion method using 24 well Linbro plate. 2 µl of 0.8 mM solution of *hCA* II in Tris-HCl pH=8.0 were mixed with of a solution of 1.5, 1.6 and 1.7 M sodium citrate, 50 mM Tris pH 8.0 and were equilibrated against 500 µl of the same solution at 296 K. Crystals of the protein grew in a few days. *hCA* II crystals were soaked in 5mM inhibitor solution for 2 days. The crystals were flash-frozen at 100K using a solution obtained by adding 25% (v/v) glycerol to the mother liquor solution as cryoprotectant. Data on crystals of the

complexes with (\pm)-**3a-3c** **MOD_23, MOD_66 and MOD_67** were collected using synchrotron radiation at the ID-11.2C beamline at Elettra (Trieste, Italy) with a wavelength of 1.000 Å and a Pilatus3_6M Dectris CCD detector. Data were integrated and scaled using the program XDS.[ref] The crystal structure of *hCA II* (PDB accession code: 4FIK) without solvent molecules and other heteroatoms was used to obtain initial phases of the structures using Refmac5.[ref]. 5% of the unique reflections were selected randomly and excluded from the refinement data set for the purpose of Rfree calculations. The initial $|F_o - F_c|$ difference electron density maps unambiguously showed the inhibitor molecules. Atomic models for inhibitors were calculated and energy minimized using the program JLigand 1.0.40. [ref].

Refinements proceeded using normal protocols of positional, anisotropic atomic displacement parameters alternating with manual building of the models using COOT.[ref] Solvent molecules were introduced automatically using the program ARP.[ref] The quality of the final models was assessed with COOT. Atomic coordinates were deposited in the Protein Data Bank (PDB accession codes: 6SX9, 6SYB, 6SYS). Graphical representations were generated with Chimera.[ref]

Crystallization and X-ray data collection of ER α and ER β complexes

ER α -Y537S LBD crystals were obtained via hanging drop vapor diffusion over a reservoir of crystallization buffer (22.5% or 25% PEG 3350, 0.2 M NaCl, 0.2 M MgCl₂, and 0.1 M HEPES, pH 7.0). To this end, drops containing 0.6 μ l of protein solution (11 mg/ml of ER α -Y537S LBD expressed and purified as previously described (PMID: **20924370**), 1 mM MOD_66/67, 1 mM NCOA2 peptide) and 0.8 μ l of the crystallization buffer were suspended over the reservoir until the crystals matured. Single-crystal X-ray diffraction data were collected at the SouthEast Regional Collaborative Access Team (SER-CAT) beamline 22-BM and **processed using HKL-2000** software.

ER α LBD crystal structures were solved by molecular replacement and automated model building using the *PHENIX* software suite (PMID: **21821126**), and **2QA8.pdb** as a starting model. Ligand coordinate files were generated using ChemOffice/Chem3D software suite (PerkinElmer). The ligands were carefully docked into vacant electron density in the ligand-binding pocket of the new models using *COOT* (PMID: **15572765**). The ligand-bound models were optimized using the *PDB_REDO* web server (PMID: **25075342**), and subject to multiple rounds of refinement, re-docking, and rebuilding as previously described (PMID: **24076406**). The structures were analyzed using *COOT* (PMID: **15572765**) and posed for presentation using *CCP4MG* (PMID: **21460457**). **Crystal structures are deposited in the PDB with accession numbers XXX and YYY**

Acknowledgements

We acknowledge the Elettra-Sincrotrone Facility for provision of synchrotron radiation facilities and we would like to thank Nicola Demitri for assistance in using beamline XRD2. We thank OpenEye Scientific Software, Inc., for a free academic license of the OpenEye Toolkit.

REFERENCES